

PATENT ABSTRACTS OF JAPAN

(11)Publication number : 04-242436

(43)Date of publication of application : 31.08.1992

(51)Int.Cl.

G06F 11/14

G06F 9/46

G06F 15/16

(21)Application number : 03-003908

(71)Applicant : TOSHIBA CORP

(22)Date of filing : 17.01.1991

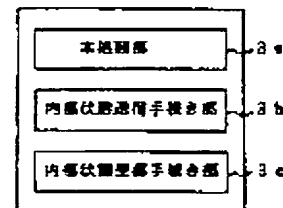
(72)Inventor : SEKI TOSHIKUNI
OYAKE YASUKUNI
TAMURA SHINSUKE

(54) STAND-BY REDUNDANT ELEMENT CONTROL SYSTEM

(57)Abstract:

PURPOSE: To copy the best internal state by obtaining a check point at the best timing individually corresponding to the processings of elements without any alteration for specifying the check point in the processings characteristic to the respective elements.

CONSTITUTION: Each program module is equipped with an internal state transmitting procedure part 3b or internal state receiving procedure part 3 which sends the internal state of its element or receives the internal states of other elements; and a main system element sends its state to a stand-by system element, which is updated according to the sent state of the main system element.



LEGAL STATUS

[Date of request for examination]

[Date of sending the examiner's decision of rejection]

[Kind of final disposal of application other than the examiner's decision of rejection or application converted registration]

[Date of final disposal for application]

[Patent number]

[Date of registration]

[Number of appeal against examiner's decision of rejection]

[Date of requesting appeal against examiner's]

decision of rejection]

[Date of extinction of right]

(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11) 特許出願公開番号

特開平4-242436

(43) 公開日 平成4年(1992)8月31日

(51) Int.Cl. ⁵	識別記号	庁内整理番号	F I	技術表示箇所
G 0 6 F 11/14		9072-5B		
9/46	3 3 0 C	8120-5B		
15/16	4 7 0 B	9190-5L		

審査請求 未請求 請求項の数1(全 4 頁)

(21) 出願番号 特願平3-3908

(22) 出願日 平成3年(1991)1月17日

(71) 出願人 000003078

株式会社東芝

神奈川県川崎市幸区堀川町72番地

(72) 発明者 関 俊文

神奈川県川崎市幸区柳町70番地 株式会社
東芝柳町工場内

(72) 発明者 岡宅 泰邦

神奈川県川崎市幸区柳町70番地 株式会社
東芝柳町工場内

(72) 発明者 田村 信介

神奈川県川崎市幸区柳町70番地 株式会社
東芝柳町工場内

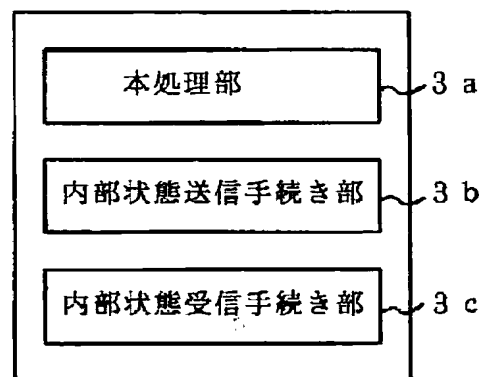
(74) 代理人 弁理士 須山 佐一 (外1名)

(54) 【発明の名称】 待機冗長要素制御方式

(57) 【要約】

【目的】 各要素固有の処理中にチェックポイントを明記するような変更をすることなく、かつ要素の処理に応じて個別に最適なタイミングでチェックポイントをとれ、最適な内部状態のコピーが可能な方式。

【構成】 各プログラムモジュールが、自要素の内部状態を送信または他要素の内部状態を受信する内部状態送信手続き部3bまたは内部状態受信手続き部3cを備え、主系要素より待機系要素に主系要素の状態を送出し、送出された主系要素の状態により待機系要素を更新するようにしたものである。



1

2

【特許請求の範囲】

【請求項1】 複数のプログラムモジュールからなるシステム中のすべてまたは一部要素を、主系要素群と待機系要素群とに多重化し、通常は主系要素間でメッセージ交換しながら与えられた作業を実行し、主系要素故障時に待機系要素が主系として処理を継続する待機冗長型システムにおいて、前記各プログラムモジュールが、自要素の内部状態を送信または他要素の内部状態を受信する第1の手段と、この第1の手段により前記主系要素より前記待機系要素に主系要素の状態を送出する第2の手段と、この第2の手段により送出された主系要素の状態により待機系要素を更新する第3の手段とを具備することを特徴とする待機冗長要素制御方式。

【発明の詳細な説明】

【0001】 【発明の目的】

【0002】

【産業上の利用分野】 本発明は、プログラムや資源のようなシステム要素が信頼性向上等の目的のため多重化され、それらシステム要素が通常時処理する主系要素群と、主系要素故障時に処理を引き継ぐ待機系要素群からなるシステムにおいて、主系要素と待機系要素の状態を同一に保つための待機冗長要素制御方式に関する。

【0003】

【従来の技術】 近年、電算機システムの各分野における要求の多様化及び技術の高度化に伴って、システムが大規模化かつ複雑化する傾向にあるため、分散してジョブを処理する並列／分散処理システムの需要が高まっている。

【0004】 このようなシステムにおいて、信頼性を向上させるために、同一の処理を複数のシステム要素で並行に実行する並列多重処理方式と、処理は単一のシステム要素で実行し実行中のシステム要素が故障した段階で待機系要素が処理を継続する待機冗長処理方式とがある。

【0005】 並列多重処理方式は、計算機故障時に待機系要素の立ち上げ等の操作がいらず、システムを停止することなく処理を継続できるが、正常時においても同一の処理を複数の多重化された要素で並列に処理するため、各計算機の処理負荷は大きい。これに対し、待機冗長処理方式では、待機系要素は実際の処理を行わず各計算機の処理負荷は小さく、計算機資源を有効に利用可能である。しかし、故障時に待機系要素が処理を引き継ぐためには、待機系要素に処理の途中経過を通報し、主系要素と待機系要素の内部状態を同一に保たなければならない。このため従来は、この通報方式としてプログラム作成者が待機系要素に途中経過を通報する箇所（チェックポイントと称す。）を明記したり、システムの持つ周期的なタイマによって定期的に主系要素から待機系要素に状態をコピーしていた。

【0006】 しかしながら、チェックポイントを明記す

る場合は、ユーザがチェックポイントを取るタイミングを意識してプログラムしなければならない。また、周期的なタイマによる場合は、要素内部の状態が変更されたか否かにかかわらず定期的にコピーされて無駄な処理をしてしまったり、より細かくチェックポイントを取りたいにもかかわらず取れないといった、システム要素個々に応じたチェックポイントを設定できないという問題がある。

【0007】

10 【発明が解決しようとする課題】 このように従来の待機冗長処理方式では、主系要素と待機系要素の内部状態を同一に保つための通報方式として、プログラム作成者がチェックポイントを明記したり、システムの持つ周期的なタイマによって定期的に主系要素から待機系要素に状態をコピーする方式が採られていた。

20 【0008】 しかし、チェックポイントを明記する場合は、ユーザがチェックポイントを取るタイミングを意識してプログラムしなければならない、また周期的なタイマによる場合は、無駄な処理をしてしまったり、より細かくチェックポイントを取りたいにもかかわらず取れないといった問題がある。

【0009】 そこで、本発明は、各要素固有の処理中にチェックポイントを明記するような変更をすることなく、かつ要素の処理に応じて個別に最適なタイミングでチェックポイントをとれ、最適な内部状態のコピーが可能な待機冗長要素制御方式を提供することを目的としている。

【0010】 【発明の構成】

【0011】

30 【課題を解決するための手段】 本発明は、システム中のすべてあるいは一部要素を、主系群と待機系群として多重化し、通常は主系要素間でメッセージ交換しながら与えられた作業を実行し、主系要素故障時に待機系要素が主系として処理を継続する待機冗長型システムにおいて、各システム要素が、自要素の内部状態を送信する手続き、ないし他要素の内部状態を受信する手続き、もしくはその両手続きを持ち、主系要素が自身の内部状態送信手続きを起動し、待機系要素の要素内部状態受信手続きに主系要素の内部状態を送る。待機系要素は主系要素の内部状態を受信すると自身の内部状態を更新し、主系要素と待機系要素間で同一状態を保つことを特徴とする待機冗長要素制御方式である。

40 【0012】 また、内部状態コピーを待機系要素側から主系要素に要求し、主系要素の内部状態送信手続きを起動し、待機系要素は主系要素から送られてくる内部状態を要素内部状態受信手続きで受信し、内部状態を更新し主系要素と同一状態を保つことも可能である。

【0013】

【作用】 本発明では、各システム要素が、上述した動作を行うようにされているので、各要素固有の処理中にチ

3

チェックポイントを明記するような変更をすることなく、かつ要素の処理に応じて個別に最適なタイミングでチェックポイントをとれ、最適な内部状態のコピーが可能である。

【0014】

【実施例】以下、本発明の実施例の詳細を図面に基づき説明する。

【0015】図2は本発明の一実施例に係るシステムの構成を示す図である。

【0016】同図において、1a、1b…はそれぞれシステム要素が登録される処理実行手段としてのプロセッサ、2a、2bはプロセッサ1a、1b…間でメッセージを交換する手段としてのメッセージ交換装置を示している。

【0017】そして、プロセッサ1a、1b…によりシステム要素の機能が実行され、システム要素間の通信はメッセージ交換装置2a、2b…により制御されるようになっている。

【0018】また、このシステムでは、信頼性を向上させる等の理由により、例えば図3に示すように、個々のシステム要素は多重化され、各プロセッサ1a、1b…に分散配置されている。

【0019】図3では、システム要素Aはプロセッサ1a、1b、1cにそれぞれ登録されているが、プロセッサ1aに登録されているシステム要素A1のみが通常時動作する主系要素であり、プロセッサ1b、1cに登録されているシステム要素A2、A3は待機している要素である。このように待機系要素が二つ以上存在する場合は、待機系要素内のいずれか一つの要素が主系になる。

【0020】システム要素Bはプロセッサ1a、1bにそれぞれ登録されており、プロセッサ1bのシステム要素B1が主系要素、プロセッサ1aのシステム要素B2が待機系要素である。

【0021】同様に、システム要素Cはプロセッサ1a、1b、1cにそれぞれ登録されており、プロセッサ1cのシステム要素C1が主系要素、プロセッサ1a、1bのシステム要素C2、C3が待機系要素である。

【0022】ここで、本実施例において、待機系要素がチェックポイントをとるタイミングで主系要素に内部状態を要求する場合について図1、図4および図5に基づいて説明する。

【0023】各システム要素は、図1に示すように、個々のシステム要素固有の処理をする本処理部3aと、待機系要素に内部状態を送信する内部状態送信手続き部3bと、主系要素から内部状態を受信する内部状態手続き部3cとを備える。

4

【0024】待機系要素例えばB2は、チェックポイントをとるべきタイミングになると、主系要素B1にチェックポイントを受け取るように要求する(ステップ401)。主系要素B1はその要求を受け取ると、内部状態送信手続き部3bを起動し、その手続きによって主系要素B1の内部状態を読み出し(ステップ501)、待機系要素B2の内部状態受信手続き部3cにその内部状態を送信する(ステップ502)。

【0025】待機系要素B2は主系要素B1にチェックポイント要求した後、主系要素B1から内部状態が送られてくるのを待つ。

【0026】主系要素B1からの内部状態を内部状態受信手続き部3cで受信すると(ステップ402)、自要素の内部状態を更新する(ステップ403)。

【0027】ここで、待機系要素B2は、主系要素B1の故障を検出すると自要素を主系とし、保存されている内部状態に基づいて処理を継続する。

【0028】従って、本実施例では、内部状態送信手続き部3b、内部状態受信手続き部3cはシステム設計者が作成できるため、必要な内部状態のみを読み出しあるいは書き込みができ、システムに応じた最適な内部状態のコピーができる。

【0029】

【発明の効果】以上説明したように、本発明よれば、各要素固有の処理中にチェックポイントを明記するような変更をすることなく、かつ要素の処理に応じて個別に最適なタイミングでチェックポイントをとれ、最適な内部状態のコピーが可能となる。

【図面の簡単な説明】

【図1】 本発明の一実施例に係るシステム要素の構成を示す図である。

【図2】 本発明の一実施例に係るシステムの構成を示す図である。

【図3】 本発明の一実施例に係る待機冗長化された要素をプロセッサに分散配置された状態を示す図である。

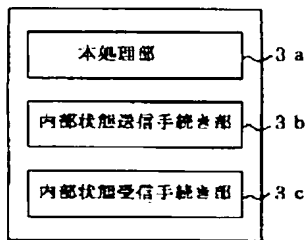
【図4】 本発明の一実施例に係る待機系要素の動作を示すフローチャートである。

【図5】 本発明の一実施例に係る主系要素の動作を示すフローチャートである。

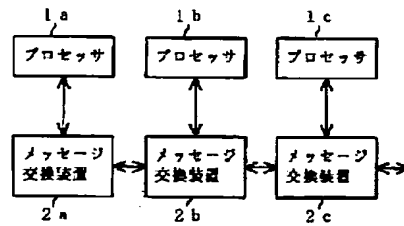
【符号の説明】

1a、1b…プロセッサ
2a、2b…メッセージ交換装置
3a …本処理部
3b …内部状態送信手続き部
3c …内部状態受信手続き部

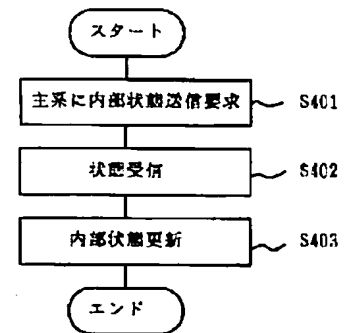
【図1】



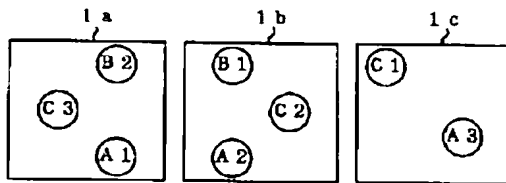
【図2】



【図4】



【図3】



【図5】

